

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

VOICE ACTIVATED, VOICE RESPONSIVE PRODUCT LOCATOR SYSTEM,
INCLUDING PRODUCT LOCATOR METHOD UTILIZING PRODUCT BAR CODE AND
AISLE-SITUATED, AISLE-IDENTIFYING BAR CODE

Inventor: Kenneth P. Glynn

Jerome R. Mahoney

Assignee: iVoice.com, Inc.

Attorney Docket No. IVC-105C

This patent application is a continuation-in-part of United States copending patent application Serial Number 09/653,658 filed on August 31, 2000 entitled "Voice Activated/ Voice Responsive Item Locator", assigned to the same assignee as designated herein and having Jerome R. Mahoney as a common inventor.

Kenneth P. Glynn, Esq.
Attorney for Applicant
Reg. No. 26,893
Glynn and Associates
24 Mine St.
Flemington, NJ 08822-1731
tele (908) 788-0077
fax (908) 788-3999

0070101700KPG-hd

5

VOICE ACTIVATED, VOICE RESPONSIVE PRODUCT LOCATOR
SYSTEM, INCLUDING PRODUCT LOCATION METHOD
10 UTILIZING PRODUCT BAR CODE AND AISLE-SITUATED,
AISLE-IDENTIFYING BAR CODE

(Attorney Docket No: IVC-105C)

15

REFERENCES TO RELATED APPLICATIONS

20

This patent application is a continuation
-in-part of United States copending patent

25

application Serial Number 09/653,658 filed on
August 31, 2000 and entitled "Voice

30

Activated/Voice Responsive Item Locator",

35

assigned to the same assignee as designated

40

herein and having Jerome R. Mahoney as a common
inventor.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to voice
activated/voice responsive item locators, i.e.
item directories, which direct a user such as a
consumer or shopper, to a specific location to
view, retrieve, order, purchase or otherwise use
the information obtained in the system. Further,
the present invention includes within the
aforesaid system, a method of collecting location
data for the system which involves the use of
product bar codes and location-situated,
location-identifying bar codes. These are read
and matched and stored in the main processor of
the system to provide location information to
subsequent users. Typically, the present
invention could be used at retail stores to
locate items to be purchased. Alternatively, it

could be used at a production facility or
distribution facility having a large number of
parts, to locate specific parts for as needed. In
other embodiments, it could be used in non-
5 commercial entities, such as public libraries to
locate a particular book. The locator of the
present invention relies upon a specific software
module to accomplish voice recognition and
response, and includes manager programming for
10 customization, updates and modifications.

2. Information Disclosure Statement

The state of the art for acquiring product
location information involves the use of manually
collected, inputted data. Bar codes have been
15 used for years to identify products, but not to
identify locations.

The following prior art patents represent various inventions relating to machines involving speech recognition for voice-based operation and thus illustrate known voice recognition applications:

U.S. Patent No. 5,111,501 to Masanobu Shimanuki describes a telephone terminal device equipped with a transmitter microphone, a receiver, a speech recognition unit that receives and recognizes speech signals from the transmitter microphone and a circuit to reduce the level of signals send from a telephone network to the receiver when the speech recognition unit receives speech signals from the transmitter microphone. Further, this device is preferably equipped with a speech reproduction

unit that reproduces the speech information
stored in a memory, in response to the
information of recognition result from the speech
recognition unit, and a circuit that prevents
5 transmission of signals from the telephone
network to the receiver when the regenerated
speech information is sent to the receiver.
Furthermore, it is desirable for this device to
be provided with a circuit that prevents
10 generation of ringing tones when an incoming call
arrives.

U.S. Patent No. 5,136,634 to David C. Rae et
al. describes voice operated facsimile machine
network which includes a method and apparatus for
15 transmitting specifically requested graphic
and/or textual data from an unattended database

stored in a memory, in response to the
information of recognition result from the speech
recognition unit, and a circuit that prevents
transmission of signals from the telephone
5 network to the receiver when the regenerated
speech information is sent to the receiver.
Furthermore, it is desirable for this device to
be provided with a circuit that prevents
generation of ringing tones when an incoming call
10 arrives.

U.S. Patent No. 5,136,634 to David C. Rae et
al. describes voice operated facsimile machine
network which includes a method and apparatus for
transmitting specifically requested graphic
15 and/or textual data from an unattended database
storage location to a requestor's facsimile

machine over a telephone line which includes a
host computer such as a PC modified with a
facsimile transmission board and a voice
generation board. The host computer receives
5 incoming phone calls and prompts the caller using
the voice board to select data files by using the
DTMF keys of a standard telephone handset. The
PC can be left unattended and can run
automatically in the facsimile transmission mode.
10 Callers can immediately access needed textual and
image data with the use of just a standard
telephone and facsimile machine. Multiple
workstation nodes can be configured in a network
setup to handle a high volume of calls in real
15 time and to allow multiple data services to
operate simultaneously.

U.S. Patent No. 5,165,095 to Mark A.

Borcherding describes a method for dialing a telephone, using voice recognition to initiate the dialing and to determine the correct telephone number. The dialing is initiated with a spoken dial command that is recognized by using speaker independent templates that are stored locally with respect to the caller's telephone. The correct telephone number is recognized by using speaker dependent template that are downloaded from a central database or by using speaker independent templates stored locally.

U.S. Patent No. 5,168,548 to Steven Kaufman et al. describes a reporting system which is disclosed herein, a speech recognizer which is used to select selections of text from a report

telephone mounted on a vehicle or similar mobile
body and which allows a call to be originated
with ease. When the user of the telephone enters
a voice command on voice inputting section, the
5 dialing unit originates a call automatically and
thereby connects the other party to the telephone
line. In a call origination procedure, the
operations for call origination and the
verifications are performed between the user and
10 the unit in an interactive sequence. In a
preferred embodiment, the unit has a particular
call origination procedure in which, when the
other party recognized by the unit is wrong as
determined by the user by verification, lower
15 place candidates for the other party are called
up in response to a particular voice command. In

an alternative embodiment, the unit indicates the other party by voicing a name for verification purpose. The alternative embodiment selects and stores only the name of the other party in response to an entered voice signal and, in the event of response for verification, combines the name having been stored and response information stored beforehand to produce composite response voice.

10

U.S. Patent No. 5,231,670 to Richard S.

Goldhor et al. describes a system and method for generating text from a voice input that divides the processing of each speech event into a dictation event and a text event. Each dictation event handles the processing of data relating to the input into the system, and each text event

15

deals with the generation of text from the
inputted voice signals. In order to easily
distinguish the dictation events from each other
and text events from each other the system and
5 method creates a data structure for storing
certain information relating to each individual
event. Such data structures enable the system
and method to process both simple spoken words as
well as spoken commands and to provide the
10 necessary text generation in response to the
spoken words or to execute an appropriate
function in response to a command. Speech
recognition includes the ability to distinguish
between dictation text and commands.

15 U.S. Patent No. 5,239,586 to Kuniyoshi Marui
describes a voice recognition system which

comprises a handset and a hands-free microphone
for generating an input audio signal, a high-pass
filter, for eliminating low frequency components
from the signal from the handset or hands-free
5 microphone, a signal level controller for
adjusting the level of the high-pass signal in
response to the user of either the handset or
hands-free microphone, a storer for storing the
speech data and a controller for controlling the
10 storer so that a user's utterance is stored or
the user's utterance is recognized by comparing
the utterance to speech data already stored. The
handset hook switch provides an on-hook control
signal to reduce amplifier gain during hands-free
15 microphone operation.

U.S. Patent No. 5,301,227 to Shoichi Kamei

et al. describes an automatic dial telephone that is useable in a motor vehicle, when a voice input is provided during a period in which input of the names of called parties is awaited, a voice pattern of the name of the called party is compared with reference patterns of called parties stored in reference patterns storing device, to determine the degree of the similarity therebetween. The names of the called parties are output to a user in the order of decreasing degree of similarity. Each time the name of a called party is output, a command word for confirmation is awaited from a user for a predetermined time period. When a voice confirmation command is input and is recognized during this waiting period, a telephone number

corresponding to the name of the called party is supplied to a channel. Consequently, the command word for confirmation may be input only if the name of the called party outputted is one desired by the user. Sensors continually monitor the driving condition of the motor vehicle in which the telephone is installed. When the operation of the steering wheel or brakes of the motor vehicle exceeds a predetermined threshold or the speed of the motor vehicle is excessive, the sensors generate safety signals that inhibit the operation of the telephone.

U.S. Patent No. 5,335,276 to E. Earle

Thompson et al. describes a communication system which is provided with multiple purpose personal communication devices. Each communication device

includes a touch-sensitive visual display to
communicate text and graphic information to and
from the user and for operating the communication
device. Voice activation and voice control
5 capabilities are included within communication
devices to perform the same functions as the
touch-sensitive visual display. The
communication device includes a built-in modem,
audio input and output, telephone jacks and
10 wireless communication. A plurality of
application modules are used with personal
communication devices to perform a wide variety
of communication functions such as information
retrievable, on-line data base services,
15 electronic and voice mail. Communication devices
and application modules cooperate to allow

integrating multiple functions such as real time
communication, information storage and
processing, specialized information services, and
remote control of other equipment into an
5 intuitively user friendly apparatus. The system
includes both desktop and hand-held communication
devices with the same full range of communication
capabilities provided in each type of
communication device.

10

U.S. Patent No. 5,349,636 to Roberto

Irribarren describes a communication system for
verbal telephonic communication which has a voice
message system for storing and retrieving voice
messages integrated with a computer database
15 accessing system for storing and retrieving text
messages from a separate computer system and for

converting the text messages into voice. The
systems are integrated via a network which
coordinates the functions of each individual
system. Additionally, the input/output ports of
5 the voice message system and the computer
database accessing system are connected in a
parallel fashion to at least one telephone line.
In this configuration a user may access both
voice messages and database information,
10 including text or electronic mail messages, with
a single telephone call. Optionally, facsimile
messages can be stored, retrieved and manipulated
with a single telephone call.

U.S. Patent No. 5,406,618 to Stephen B.
15 Knuth et al. describes a telephone answering
device that is activated by a proximity sensor

when a user crosses its field of detection and
whose operation is controlled by simple voice
commands. The device incorporates speaker-
independent voice recognition circuitry to
5 respond to spoken commands of the user that are
elicited by a system generated voice request
menu. The telephone answering device performs
all the basic functions of a telephone answering
machine in response to these simple commands and
10 there is no need for the user to manually operate
the telephone answering device.

U.S. Patent No. 5,602,963 to W. Michael
Bissonnette et al. describes a small, portable,
hand-held electronic personal organizer which
15 performs voice recognition on words spoken by a
user to input data into the organizer and records

voice messages from the user. The spoken words
and the voice messages are input via a
microphone. The voice messages are compressed
before being converted into digital signals for
5 storage. The stored digital voice messages are
reconverted into analog signals and then expanded
for reproduction using a speaker. The organizer
is capable of a number of different functions,
including voice training, memo record, reminder,
10 manual reminder, timer setting, message review,
waiting message, calendar, phone group select,
number retrieval, add phone number, security and
"no" logic. During such various functions, data
is principally entered by voice and occasionally
15 through use of a limited keypad, and voice
recordings are made and played back as

appropriate. A visual display provides feedback
to the user. During the various function, the
user can edit various different data within the
organizer by eliminating or correcting such data
or entering new data.

5

U.S. Patent No. 5,621,658 to Brion K.

Jackson describes an action contained within an
electronic mail object which is communicated from
a data processing system to another data
processing system via an audio device. The
action is executable on a data processing system.
At the sending data processing system, the action
is converted to a predetermined audio pattern.
The electronic mail object may contain text in
addition to an action. The text is also
converted to an audio pattern. The audio

10

15

patterns are then communicated to the audio
device over telephone lines or other
communication medium. At the receiving end, the
audio device records the object. A user can
5 provide the recorded object to a data processing
system, which then executes the action and
converts the text audio patterns back to text.
In addition, the action can be converted to text
and displayed on the data processing system.

10 U.S. Patent No. 5,631,745 to John J. Wong et
al. describes a telephone terminal adapted for
business or home use that includes the ability to
receive and send facsimiles, a voice answering
function and a computer modem. Various input and
15 output devices may be used for the facsimile
function. A voice annotated facsimile may be

sent and received. At the same time the
facsimile is viewed on a video monitor or
ordinary television set, an accompanying voice
message is heard through the sound system of the
5 monitor or television set. The terminal has an
architecture including a central processor and an
internal bus structure to which several types of
memory, various input-output devices and an
interface with the telephone line are connected,
10 among others. Audio Random Access Memory (ARAM)
is used for storing both facsimile data and voice
data.

U.S. Patent No. 5,671,328 to Gregory P.
Fitzpatrick et al. describes a method and data
15 processing system which are disclosed for
automatically creating voice processing template

entries. In one embodiment, the invention automatically assembles a plurality of commands received by the data processing system, at least one of said commands having a voice recognition criteria component associated therewith, counts the occurrences of the plurality of commands, assembles voice recognition criteria components associated with the plurality of commands, and, as a result of the occurrence count exceeding a predefined minimum, constructs a voice recognition template entry by associating the assembled voice recognition criteria components with the assembled plurality of commands.

U.S. Patent No. 5,850,627 to Joel M. Gould et al. describes a word recognition system which can: respond to the input of a character string

from a user by limiting the words it will
recognize to words having a related, but not
necessarily the same, string; score signals
generated after a user has been prompted to
5 generate a given word against words other than
the prompted word to determine if the signal
should be used to train the prompted word; vary
the number of signals a user is prompted to
generate to train a given word as a function of
10 how well the training signals score against each
other or prior models for the prompted word;
create a new acoustic model of a phrase by
concatenating prior acoustic models of the words
in the phrase; obtain information from another
15 program running on the same computer, such as its
commands or the context of text being entered

into it, and use that information to vary which words it can recognize; determine which program unit, such as an application program or dialog box, currently has input focus on its computer and create a vocabulary state associated with that program unit into which vocabulary words which will be made active when that program group has the focus can be put; detect the available computational resources and alter the instructions it executes in response; test if its ability to respond to voice input has been shut off without user confirmation, and, if so, turn that ability back on and prompt the user to confirm if that ability is to be turned off; store both a first and a second set of models for individual vocabulary words and enable a user to

selectively cause the recognizer to disregard the
second set of models for a selected word; and/or
score a signal representing a given word against
models for that word from different word model
sets to select which model should be used for
future recognition.

Notwithstanding the prior art, the present
invention is neither taught nor rendered obvious
thereby.

SUMMARY OF THE INVENTION

A voice activated/voice responsive item
locator system is disclosed to enable a user to
speak into the system and have the system respond
with location information for an item requested
by the user. For example, shopper at a home
supply store may pick up a locator phone or just

5 speak into a wall mounted or otherwise situated
microphone and say "Locate Outdoor Paint" or
"Find Hammers" or simply state what is sought
without the use of a verb, e.g. "Caulking". The
system may reply either with voice or visual
(words on a screen, or map), or both voice and
visual, e.g. "Aisle 3, Shelf 4". In some
instances the system will reply, for example,
with a "Repeat", or "Restate in different words"
10 or "Please talk to information desk" or other
default instructions.

The present invention also includes a
method of creating data for locating items so
that the system is efficiently loaded with
15 location data both prior to use by the customers
or other users, as well as so that the system may

be updated as desired while it is in use. This method involves utilization of bar codes to determine item identity, and the use of separate bar codes to determine locations. These separate bar codes are physically located on location structure, e. g. on aisle ends, shelf edges, bin walls, parking spaces, etc. This location data is read in conjunction with item identification data by bar code readers, fed to a processor in a recognizable combined format, and then stored and used as the resource data of the locator system.

For example, a supermarket could assign unique bar codes to each aisle, create bar code labels and attach them to the ends of each aisle, and then program the system according to the following simple process:

a) The processor will be programmed to read and identify products by the universal price code ("UPC") inputs from a bar code reader, and will likewise be programmed to recognize and identify locations by bar code inputs from a bar code reader, that is, the processor will be programmed to understand the codes created for particular locations to be included in the supermarket product location system;

b) The processor will also be programmed to match items (products) to locations when read between identical location readings. In other words, when a reader inputs a location bar code from one end of an aisle, and then reads all of the UPCs of all items in the aisle, and then reads the same location bar code at the other end

of the same aisle, this tells the processor to
create a matching set of pairs of products and
locations for all products read between each end
of that aisle. In an alternative embodiment, each
5 type of item could be read before or after the
location reading to create location data
pairings.

The overall locator system may be a stand
alone device, but in most embodiments would be
10 part of an internal connected system. It could be
an intranet or secured internet system, but would
in many cases be a storewide system with a
plurality of user locations (units, phones, or
microphones, with feedback at each location). The
15 system will include an embedded voice-driven
interface for speech control of: (1) operational

instructions; (2) core system locator function
operations, that is, recognition of specific
requests and responses thereto; and, (3) optional
and default functions. In preferred embodiments,
5 the present invention device is both operated by
speech (speech or voice activated) and speech
responsive (voice answers and instructions to the
user from the system). Thus, the present
invention device relies upon automatic speech
10 recognition (ASR), either in place of or in
addition to manual locator systems, e.g. book,
list, map and computer directories. In some
embodiments, user feedback features are included
wherein both audio and visual feedback is given
15 to a user in response to recognizable voice
signals, while in other possible embodiments, the

user may designate audio or visual.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention should be more fully understood when the specification herein is taken in conjunction with the drawings appended hereto wherein:

Figures 1a and 1b show a general schematic diagram showing software and functional features of a present invention item locator system, including the method of creating item /location data pairs;

Figure 2 shows a schematic diagram illustrating the physical functions of a present invention voice recognition item locator device after the item/location information data pairs have been created; and,

Figure 3 shows a schematic diagram of a present invention device illustrating details of a voice recognition submodule used therein.

DETAILED DESCRIPTION OF THE PRESENT INVENTION

5 The present invention is a voice

activated/voice responsive item locator and

system. By "item" is meant a place or thing that

a user desires to locate. Thus, a item could be a

particular brand of canned string beans, a type

10 of outdoor stain, a booth at a convention, a

particular part in inventory for sale, assemblage

or distribution, a particular automobile in a

production facility lot or in a large parking

garage, or a room, a functional group or a person

15 in an office building or the like. The response

may be in the form of a word or sentence

presented visually or audibly and it may designate an aisle, a shelf, a bin number, a row number, a row and slot or space, etc.

5 The voice recognition system digitizes words spoken via a receiver (microphone) handset, headset, or built-in microphone for conversion from analog to digital utilizing a continuous speech recognition digital signal processor (DSP). The main support structure may be a
10 conventional type housing for phones and other communications devices, may be of a different shape or configuration or may be built into a device such as a wall or desk unit, with or without monitor. It may be portable or
15 permanently affixed and could be powered by any means available, e.g. AC or DC current. In the

portable mode, the system would be wireless for the user and would, in that respect operate like a cell phone, two way radio, "walkie talkie" or other short distance wireless device, but would have a processor at a central or fixed location having the same features as described above, i.e., the DSP with programming capabilities, etc.

The DSP is connected to a programmable microprocessor and either by customized input or a standard program, the system enables the user to quickly enter voice-activated fields, e.g., such as "Where is...", "Find...", etc.

Verification of voice recognition accuracy (prior to execution) is optional and may be accomplished via synthesized voice playback and/or a screen confirmation which requires a "YES" or "NO" to

execute or open for revision. In some preferred embodiments, a screen, e.g., LCD, enables visual feedback during input phase, with support for deletion, insertion, correction, etc.

5 Cancellation of the entire command or programming instructions may be possible at any time (prior to execution), via keystroke or voice command.

Another important aspect of the present invention is the inclusion into the system of software and hardware (equipment) to utilize a
10 method of creating item location information for the system. It involves using item-identifying bar codes on items to be included and using location-identifying bar codes from corresponding
15 locations. The location-identifying bar codes are physically situated on the locations themselves.

For example, they are located on aisle ends,
shelves, bins, drawers, floor area grids, etc.

The location-identifying bar codes may be
custom created for the locations or may be
5 established as a universal location system.

Alternatively, a manager could use existing UPC
bar codes for the locations, provided that they
were different from the items to be located, and
provided that the system were programmed to
10 correlate these particular codes to specified
locations.

The item-identifying bar codes are typically
located on the items themselves, but when more
than one identical item is included, a single
15 item of the set of identical items will be
sufficient for the method to work. However, it is

preferred that all items in each set have the bar
code located thereon. In some preferred
embodiments, the bar codes for the items are
Universal Price Code (UPC) bar codes, but the
5 present invention need not be limited thereto,
such as when it would be more appropriate to
create unique identifying codes for each and
every item, such as automobiles, artwork, etc.

The essential features of the present
10 invention involve the creation of a voice-based
guide or locator and the creation of appropriate
item/corresponding location data base, to offer
enhanced convenience and speed to users for
location of one or more items.

15 Figures 1a and 1b show a general schematic
diagram of a present invention system showing

general software features and functional
features. Thus, the present invention system
includes a method, software and hardware for the
creation of item/location data pairs, as
5 described above. In Figure 1a, the basic aspects
of the item/location information data creation
are set forth in schematic form. The unique
item-identifying bar codes are attached 2 to at
least one of each different item for a plurality
10 of sets of items, each set having items different
from the items in the other sets. Likewise,
unique location-identifying bar codes are
attached 4 to the corresponding locations, and,
subsequently, they are read 6 in predetermined
15 manner so that the program recognizes sequences
and creates data pairs to develop the

item/location vocabulary for the system. This information is included in manager inputs 10 (reference also Figure 1b). The method shown in Figure 1a is repeated as needed for updating 8.

5 Figure 1b illustrates other features of the present invention and includes a central processor 1 which may be an external or internal component, i.e., within a single unit or at a separate location from audio receivers and
10 transmitters , e.g., microphones/speakers for user inputs and feedback to users.

 The system may be preprogrammed with the user being required to follow concise instructions for activation and operation, or may
15 be programmable to alter, add or enhance ease or methods of use, e.g. through a limited access

code, for manager inputs 3 of user instructions.

In any event, manager inputs 3 shall include

functional selections and inputs of items and

their locations, with provision for subsequent

5 access for modifications. This programming may

include direct keyboard, voice, etc., and, as

mentioned, may include security capabilities for

preventing unauthorized use, e.g. voice

identification (user recognition) or user

10 security code system, as well as other options

which may be included therein, such as a "help"

detailed manager instruction section.

Once the system has been programmed for

use, the user operation unit(s) 5 provide

15 functional access, which may be passive, i.e.,

the user speaks, picks up a phone, presses a

button, or otherwise takes some action to
activate the system; or it may be active, i.e., a
proximity sensor, a periodicity timer, or other
internal mechanism may automatically activate the
5 system and could trigger an audio or visual
query, such as "May I help you locate a product?"

Once the system has been activated and a
user has stated the necessary words of input to
activate the device, recognition/non-recognition
10 response 7 results from processing the user
inputs to central processor 1 , and audio and/or
video response unit(s) 9 provide feedback 11 to
the user, either by answering the inquiry,
conditionally defaulting, e.g., asking for a
15 repeat or a restate the question, or fully
defaulting, e.g. directing the user to a courtesy

desk or check out counter for help.

Figure 2 shows a schematic diagram illustrating a present invention voice activated/voice responsive item locator system, showing the physical arrangement and function of components after the item/corresponding location information has been inputted. Thus, symbol 17 indicates an optional user prompter proximity sensor and symbol 21 is a microphone or equivalent component for voice input. The voice input is sent to audio controller 19 and to automatic speech recognition unit 23 and is converted from analog to digital signals. CPU/Memory 25 compares the digital signals to the set up or dictionary of digital words or phrases in memory. Once a match is found, the system

processor 27 and data storage 31 operate to respond with an answer or a default instruction or a query by providing digital text to text-to-speech generator 29, which provides audio feedback to a user via audio controller 19 and speaker 33. Feedback to a user may also be provided on visual screen 37 via display controller 35. Keyboard 39 is used for manager set up and modifications.

Figure 3 shows the details of one preferred embodiment of the submodule used in the present invention device. The voice recognition component converts an acoustic signal into a sequence of labels. The system takes the raw acoustic data, and processes it through the recognizer. The recognizer then matches it

against a set of models using a decoder that
generates a recognition token. This token
represents what the user said as either a single
word or utterance. The recognizer itself does
5 not interpret the meaning of the recognized
output, that is the function of the interpreter
(described later). The recognizer uses Hidden
Markov Models (HMMs) to provide for a continuous
speech recognition engine. HMMs do not process
10 the acoustic signal directly but instead split
the signal into a sequence of discrete
observations. These observations are derived
from a digital representation of the signal that
had been converted from the analog signal
15 generated by the microphone. During recognition,
the likelihood of each model (or sequence of

models) matching the incoming signal is
calculated. The recognizer simply selects the
most likely model to decode the signal. As this
is done continuously, the recognizer can process
5 speech as opposed to isolated words, allowing the
user to talk more naturally.

Each acoustic model represents a short
sound. The interpreter combines these sounds
into words using a dictionary. This dictionary
10 specifies the pronunciation of each word in terms
of the acoustic models. After identifying the
most likely word, the interpreter then joins sets
of models together (using a Viterbi decoder) in a
series of pre-defined connections such that paths
15 can be established to provide for a degree of
"natural language" recognition; in other words,

the user can say "Find hammers", "Where are hammers" or "hammers" and they are all understood to mean the same thing. Moreover, these sets of models and dictionaries are interchangeable, allowing the same voice recognition component to be used in a variety of applications.

As the voice recognition component is running continuously, there needs to be a way to distinguish background conversations that might accidentally trigger an unwanted action by the device. For example, two people standing by a voice-activated device might be discussing locations of different goods in a supermarket and be misinterpreted or undesireably responded to. To avoid this problem, the recognition unit requires a command word to trigger before

beginning further recognition. The trigger word is a user-definable setting.

Thus, in Figure 3, initialization 51 initiates monitoring 53 for a trigger word from a user. When a word is received, it is analyzed to determine whether or not a trigger word 55 has been received. If not, signal 57 returns the status to monitoring 53 for a new word. This loop continues until a trigger word is recognized and an inactivity timer 59 is started. The monitor 61 proceeds with the monitoring for the next word and waits for timer pop 65. When an event 63 is received, timer pop 65 returns to the monitor 53 to continue the monitoring process and the voice data is sent to interpretation 67. If it is understood 69, an action 75 is processed and

feedback function 77 is performed. Additionally,
signal 79 prompts user 71. Likewise, if the
interpretation is not understood 69, user 71 is
prompted and via signal 73, timer 59 begins
again. These cyclings operate on a continual
basis while the system is initiated. Voice
activation may also be used to shut down the
system.

Obviously, numerous modifications and
variations of the present invention are possible
in light of the above teachings. It is therefore
understood that within the scope of the appended
claims, the invention may be practiced otherwise
than as specifically described herein.